Ongoing work on MRA of Traffic Matrices

David Rincón[1], Matthew Roughan[2]
[1]Universitat Politècnica de Catalunya, [2]University of Adelaide
Email: drincon@entel.upc.edu, matthew.roughan@adelaide.edu.au

I. Introduction

Internet Traffic Matrices (TMs), giving traffic volumes from ingress to egress nodes in a network, are a basic input to many network engineering tasks, but are non-trivial to measure, and so much work has gone into measurement or their indirect inference [1] from readily available link load measurements. In the latter case, however, the measurements do not provide enough information to form a well-posed inference problem, and therefore some type of side information (usually in the form of a traffic matrix model) is needed to perform the inference. Network engineering with TMs also needs models. For instance, we may need to *predict* a traffic matrix at some time in the future. In addition, we may wish to detect anomalous traffic behaviour, and a simple approach is to look for large deviations from predicted behaviour.

The problem of finding a "good" model for TMs is problem dependent. The criterion we focus on here is that the traffic matrix model should be *sparse*. A traffic matrix for a network with N nodes has N^2 terms, and since N can be in the thousands, the number of terms in the traffic matrix can become very large. A sparse model has a number of parameters $M \ll N^2$. There are good reasons for a sparse model:

- There is a tradeoff between model fidelity and the model's predictive power. A model with a large number of parameters may work well for one dataset, but provide bad predictions because it is too specific.
- If the model has few parameters then we have more hope of attaching physical meaning to them.
- The inference problem is ill-posed when we have $K = O(N)$ link-load measurements, but N^2 parameters that we need to estimate. If the TM model had $M \leq K$ parameters, then the problem might become well posed.

Often, approaches to TM inference have sought some kind of sparse model for the matrices, with a view that this will bring the problem back towards being well-posed. The gravity model [1] is a good example, with only $2N$ parameters. However, in this case the model itself is not a particularly accurate representation of a TM, it simply forms a *prior* used in a regularization approach for inference.

Our approach for finding such a sparse model is to use Multi-Resolution Analysis (MRA). The sparseness of the coefficients is the leading reason why modern audio and video compression techniques often use MRA techniques such as wavelets. Unfortunately, standard wavelet-based MRA analysis is not appropriate for traffic matrices. In a TM the spatial relationships between elements are more complicated than in an image, which is a rectangular grid sampling of a two-dimensional field. We exploit the new Diffusion Wavelets (DW) approach [2] and perform multi-resolution analysis of functions defined on graphs. The graph may represent the underlying network (over which our traffic matrix is routed), and thus reflect the natural spatial relationships in the TM (i.e., two traffic matrix elements originating from locations close together in the network may share characteristics such as their diurnal traffic pattern).

One of our contributions is to generalize Diffusion Wavelets to 2D and apply them to modelling traffic matrices, which we see as two-dimensional functions of the nodes. In our first experiments with real data from operational networks we found that in the DW domain traffic matrices are sparse.

II. MRA, Wavelets & Diffusion Wavelets

The Discrete Wavelet Transform (DWT) analyzes signals by computing its scalar product with dilated (by powers of 2) and translated versions of the *mother wavelet* function, thus analyzing the input signal at time scales $t = 2^j$. In mathematical terms, we obtain a MRA with a set of nested approximation (*scaling*) subspaces V_j, $V_1 \supset V_2 \supset \ldots \supset V_J$ and their orthogonal complements, the high-frequency detail (*wavelet*) subspaces $W_j = V_{j-1} - V_j$.

The aforementioned *classical* wavelet transforms operate on signals defined on uniformly sampled grids on R and R^2, respectively. However, a TM is not defined on a regular lattice — it is defined across a computer network, which can be represented by a graph. Diffusion Wavelets [2] are a generalization of the wavelet transform in which the MRA can be performed on structures such as manifolds or graphs. In our case the underlying structure is a graph $G\{V, E\}$ (where V and E are the vertex and edges sets, respectively). We wish to analyze a function $f : V \to R$, i.e., we have a function $f(i)$, which maps each vertex i to a real number.

The approach is to create a *diffusion operator* that plays the role of the dilation in the DWT. Application of the diffusion operator "blurs" the original function, but in a way that is adapted to the underlying graph. Locations that are close together in the graph will be blurred into each other, while locations that are far apart will remain separated. Mathematically we represent the diffusion operator by a linear transform Tf for which there are many choices. Simple examples include a heat-like diffusion across the graph, or a stochastic matrix representing a random walk on the matrix. The dilation consists in take powers of the matrix T. Intuitively, if a diffusion continues over n time steps, we would apply the linear transform n times, i.e., $T^n f$. This results in successive blurring of the function, as required. In the random walk interpretation, assume f represents an initial distribution of states, then $T^n f$ represents the state distribution after n time steps, which we know will tend to blur towards the equilibrium distribution. For graphs, the natural equivalent to the frequency-based decomposition resulting from the DWT is spectral graph theory, i.e., the study of the eigenvalues and eigenfunctions of linear operators [3]. The eigenvalues of T^n are λ_i^n, and the eigenvectors remain invariant with respect to n. As $n \to \infty$ all of the eigenvalues $|\lambda_i| < 1$ will tend to zero, and eventually they will fall below the threshold ϵ. As such, the successive application of the (now approximated) diffusion operator will break the graph spectrum into subbands, much as the classical wavelet transform does.

Traffic matrices can be represented as two-dimensional functions $F(v_1, v_2)$ of pairs of vertices where v_1 is the ingress node, v_2 is the egress node, and $F(v_1, v_2)$ is the traffic volume from v_1 to v_2. Hence, we need to extend DWs to 2D. In [4] we presented a 2D DW by computing the tensor product of the 1D DW bases, and proved that the 2D DW transform is invertible and orthonormal.

III. DW-based MRA of traffic matrices

In our first experiments with the 2D DW tool we have studied over 20000 traffic matrices belonging to two datasets from the Abilene and GÉANT networks, with 12 and 23 PoPs, respectively. The granularity of the TMs is 5 minutes in Abilene and 15 minutes for GÉANT. For more details about the datasets refer to [5]. The TMs were analyzed with the 2D Diffusion Wavelet transform, with two goals in mind. First, we wanted to visualize how the diffusion process affected a traffic matrix, in order to develop our intuition about the multi-resolution decomposition and check the invertibility and perfect reconstruction properties. This results are reported in [4]. Secondly, we wanted to assess the compressibility (sparsity) obtained with the 2D DW. We performed several tests on fortnight-long and month-long series from both datasets in order to assess the extent to which the energy of the original matrices is compressed in a few coefficients. Figure 1 shows the Mean Square Error (MSE) of the reconstructed TMs versus the percentage of coefficients used in the reconstruction of two representative, month-long traces, studied with the adjacency/random-walk operator. The results confirm the sparseness of the DW representation: on average, 15% of the coefficients retain more than 90% of the original TM energy.

We are now experimenting with alternative operators. For example, we have tested a gravity-model-based operator in which the diffusion operator takes the form of the gravity model (i.e., the rank-1 approximation) of the TM. Recall that this data can be easily obtained from SNMP counters. The achieved compressibility is considerably higher, which is not a surprise, since the operator is somehow extracting the intrinsic correlation preserved in the low-rank representation of the TM.

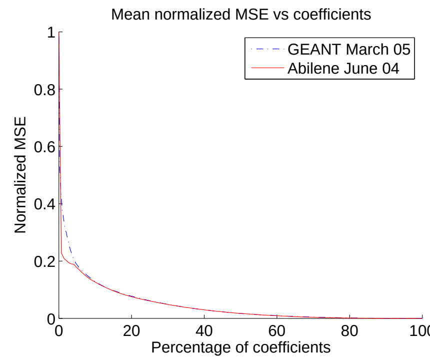

Fig. 1. Mean normalized MSE vs percentage of coefficients for two month-long TM traces, analyzed with the topology/random walk operator.

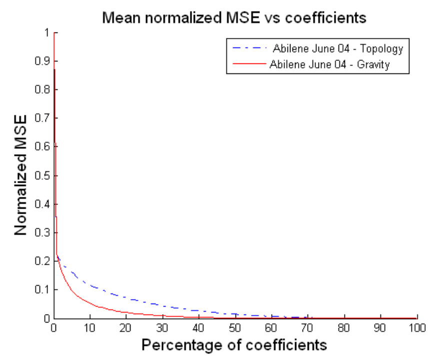

Fig. 2. Comparison of MSE for the topology/random walk and the gravity-model-based operators for a month-long Abilene TM trace.

Figure 2 shows the comparison between the topology/random-walk and gravity model-based operators. We have also tried with a routing-based operator and a random walk-plus-time-correlation 3D operator, but the (non-conclusive) results are not as good as with the aforementioned operators.

Our results are presented as a "proof of concept", a first step towards creating viable sparse models of TMs for use in the various tasks mentioned above: *inference, synthesis, and prediction*.

IV. Acknowledgments

D. Rincón was supported partially by UPC and by EuroNF . M. Roughan was supported in part by ARC grant DP0665427.

References

[1] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg. Fast accurate computation of large-scale IP traffic matrices from link loads. In *ACM SIGMETRICS*, pages 206–217, San Diego, California, June 2003.

[2] R. R. Coifman and M. Maggioni. Diffusion Wavelets. *Applied and Computational Harmonic Analysis*, 21(1):53–94, July 2006.

[3] F. Chung. *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)*. American Mathematical Society, February 1997.

[4] D. Rincon, M. Roughan, and W. Willinger. Towards a meaningful MRA of traffic matrices. In *ACM Internet Measurement Conference*, pages 331–336, October 2008.

[5] S. Uhlig, B. Quoitin, J. Lepropre, and S. Balon. Providing public intradomain traffic matrices to the research community. *SIGCOMM Computer Communication Review*, 36(1):83–86, 2006.